NOVA IMS Information Management School

07
Sentiment Analysis
Rule Based Approach

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

AGENDA

1. What is Sentiment Analysis?
2. What are rule-based approached algorithms?
3. Use case

1

What is sentiment analysis?

The process of detecting positive or negative sentiment in text

What is sentiment analysis

Introduction

What is sentiment analysis?

*Sentiment analysis uses natural language processing (NLP), text analysis, and computational techniques to automate the **extraction or classification of sentiment** from sentiment reviews.*

Hussein, D. M. E. D. M. (2018). A survey on sentiment analysis challenges. Journal of King Saud University-Engineering Sciences, 30(4), 330-338.

*Sentiment analysis or opinion mining is the computational **study of people's opinions, sentiments, emotions, appraisals, and attitudes towards entities** such as products, services, organizations, individuals, issues, events, topics, and their attributes*

Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1253.

In sentiment analysis we are focused on the **polarity** of a text:

What is sentiment Polarity?

- A numeric score given to assess both the favorable and unfavorable elements within a text document.
- This rating is determined by **subjective criteria** such as particular words and expressions that convey emotions and sentiments.

We can go further and try to detect specific feeling or emotions:

The Scherer typology of affective states

Emotion: Relatively brief episode of response to the evaluation of an external or internal event as being of major significance. (angry, sad, joyful, fearful, ashamed, proud, elated, desperate)

Mood: Diffuse affect state, most pronounced as change in subjective feeling, of low intensity but relatively long duration, often without apparent cause. (cheerful, gloomy, irritable, listless, depressed, buoyant)

Interpersonal stance: Affective stance taken toward another person in a specific interaction, coloring the interpersonal exchange in that situation. (distant, cold, warm, supportive, contemptuous, friendly)

Attitude: Relatively enduring, affectively colored beliefs, preferences, and predispositions towards objects or persons. (liking, loving, hating, valuing, desiring)

Personality traits: Emotionally laden, stable personality dispositions and behavior tendencies, typical for a person. (nervous, anxious, reckless, morose, hostile, jealous)

Scherer, K. R. (2000). Psychological models of emotion. Borod, J. C. (Ed.), The neuropsychology of emotion, 137–162. Oxford.

The Scherer typology of affective states

Emotion: Relatively brief episode of response to the evaluation of an external or internal event as being of major significance. (angry, sad, joyful, fearful, ashamed, proud, elated, desperate)

Mood: Diffuse affect state, most pronounced as change in subjective feeling, of low intensity but relatively long duration, often without apparent cause. (cheerful, gloomy, irritable, listless, depressed, buoyant)

Interpersonal stance: Affective stance taken toward another person in a specific interaction, coloring the interpersonal exchange in that situation. (distant, cold, warm, supportive, contemptuous, friendly)

Attitude: Relatively enduring, affectively colored beliefs, preferences, and predispositions towards objects or persons. (liking, loving, hating, valuing, desiring)

Personality traits: Emotionally laden, stable personality dispositions and behavior tendencies, typical for a person. (nervous, anxious, reckless, morose, hostile, jealous)

Scherer, K. R. (2000). Psychological models of emotion. Borod, J. C. (Ed.), The neuropsychology of emotion, 137–162. Oxford.

1 What is sentiment analysis
Introduction

Sentiment analysis is the detection of attitudes

“enduring, affectively colored beliefs, dispositions towards objects or persons”

1. **Holder** (source) of attitude
2. **Target** (aspect) of attitude
3. **Type of attitude**
   - From a set of types: Like, love, hate, value, desire, etc.
   - Or (more commonly) simple weighted polarity: positive, negative, neutral, together with strength
4. **Text** containing the attitude
   - Sentence or entire document

Tasks in Sentiment analysis:

Simplest task:

Is the attitude of this text positive or negative?

More complex:

Rank the attitude of this text from 1 to 5

Advanced:

Detect the target, source, or complex attitude types

1.1 What is sentiment analysis

Why is sentiment analysis important?

1. **The growing significance of sentiment analysis**
    - In today's digital age, people express their thoughts and emotions more openly than ever
    - Sentiment analysis is now a ***vital tool for monitoring*** and comprehending sentiment across various types of data

2. **Sentiment analysis for Customer insights**
    - The automated examination of customer feedback, including opinions shared in survey responses and social media conversations, provides brands with invaluable insights.
    - These insights enable brands to discern ***what pleases or frustrates their customers***, facilitating the customization of products and services to align with customer preferences.

1.1 What is sentiment analysis

Why is sentiment analysis important?

3. **Uncovering Customer sentiment at every stage**
    - For instance, employing sentiment analysis to automatically assess over 4,000 open-ended responses in your customer satisfaction surveys can offer a comprehensive understanding of ***why customers feel content or dissatisfied at different points*** in their customer journey.

4. **Real-time Customer Sentiment Tracking**
    - In some cases, there is a need to ***track brand sentiment in real-time*** to swiftly identify and address displeased customers. This immediate response can mitigate potential issues and enhance customer satisfaction.

Why is sentiment analysis important?

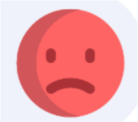
5. **Analyzing Sentiment trends over time**
   - Brands may want to compare sentiment data from one quarter to the next to determine if corrective measures are necessary.
   - Subsequently, examining deeper qualitative data can provide insights into the underlying reasons for ***changes in sentiment***, whether it is on the rise or decline.

6. **Consistent criteria**
   - Human sentiment assessment is subjective and often inconsistent (60-65% agreement).
     - Personal experiences, thoughts, and beliefs influence sentiment tagging.
   - Centralized ***sentiment analysis systems ensure consistent criteria*** across all data.
     - Enhanced accuracy and deeper insights result from standardized criteria.

1.2 What is sentiment analysis

Sentiment Analysis Challenges

There are some cases where it is easy to understand the sentiment behind:

- MSI has the best selection of laptops 
- Apple has a great design 
- I really like crime series 
- I hate waiting for the next episode to come out 

What is sentiment analysis

Sentiment Analysis Challenges

But some are more difficult to interpret…

- I *do not dislike* metal music **(phrase with negation)**
- Disliking horror movies *is not uncommon* **(negation, inverted word order)**
- *Sometimes* I really hate metal music **(The adverb sometimes modifies the sentiment)**
- I love having to wait one year for the next season to come out! **(sarcasm)**
- The final episode was unexpected with a *terrible* twist at the end **(negative term used in a positive way)**
- The movie was easy to watch but I would not recommend it **(difficult to categorize)**

!! Challenges in Sentiment Analysis !!

Context-dependent errors

Sarcasm

People tend to use sarcasm as a way of expressing their negative sentiments, but the words used are positive!

" I am delighted to have Text Mining on Monday mornings!"

Polarity

The emotional tone in some sentences can be very clear and apparent

"It was a terrible experience!"

Others are not easily classified as positive, negative or neutral

"The service quality is not mentionable"

Context-dependent errors

Polysemy

Words can have more than one meaning!

When "bat" refers to the flying mammal, it's typically a neutral term. For example, "I saw a bat in the cave."

In the context of sports, "bat" can be positive or negative. "He hit a home run with the bat" is positive, while "He missed out with the bat" is negative.

As a verb, "bat" can mean to hit something or someone. In this case, it might be negative, as in "He batted me on the head."

Emojis

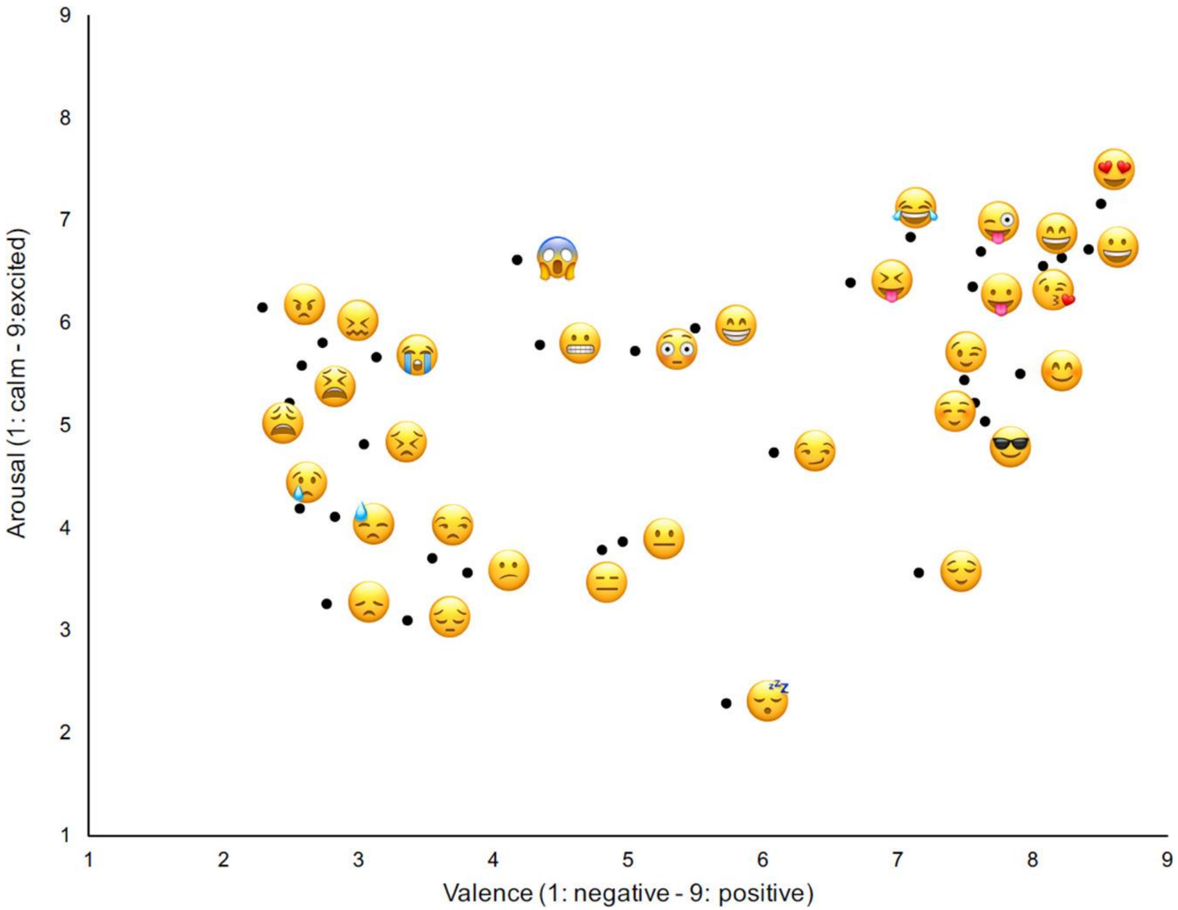


Image Source

- Emojis have become a part of daily life and are more effective in expressing one's sentiments compared to words.
- However, as the sentiment analysis tools depend on written texts, emojis sometimes cannot be classified accurately and thus are removed from many (but not all) analyses.
- If those are removed from text, one ends up sometimes with a noncomprehensive analysis.

Potential Biases in Model Training

If the algorithm associates traditionally masculine roles or traits (e.g., "strong," "leader") with positive sentiment and traditionally feminine roles or traits (e.g., "nurturing," "empathetic") with negative sentiment, it could produce biased sentiment scores.

This bias potentially leads to inaccurate and unfair assessments based on gender stereotypes.

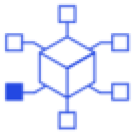
1.2 What is sentiment analysis

Sentiment Analysis Challenges

Minimizing bias will be critical if artificial intelligence is to reach its potential and increase people's trust in the systems.

Six potential ways forward for artificial-intelligence (AI) practitioners and business and policy leaders to consider

1



Be aware of contexts in which AI can help correct for bias and those in which there is high risk for AI to exacerbate bias

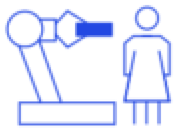
2

Establish processes and practices to test for and mitigate bias in AI systems

3

Engage in fact-based conversations about potential biases in human decisions

4



Fully explore how humans and machines can best work together

5

Invest more in bias research, make more data available for research (while respecting privacy), and adopt a multidisciplinary approach

6

Invest more in diversifying the AI field itself

There is still a long way for artificial intelligence to achieve its maximum potential!

McKinsey & Company

1.3 What is sentiment analysis

How does sentiment Analysis work?

Sentiment Analysis Algorithms

Rule-Based

Based on a set of manually crafted rules

Automatic

Rely on Machine Learning to learn from data (SVM, NN, NB, …)

Hybrid

Combine rule-based with automatic approaches

1.4 What is sentiment analysis

Rule-based vs Machine Learning

Rule-based Approaches Advantages

1. **Accuracy**

- Work based on specific rules. They follow these rules strictly to make sure they are precise and accurate in their tasks.

2. **Ease of use**

- Don't need a lot of complex data. They're straightforward and simple to create, use, and fix. This makes them developer-friendly.

3. **Speed**

- Once trained properly, rule-based systems make decisions fast. They don't need to think or learn; they follow their rules, which means they can respond quickly.

1.4 What is sentiment analysis
Rule-based vs Machine Learning

Rule-based Approaches disadvantages

1. **Limited in what they can do**
   - Rule-based systems are very precise but inflexible. They can't learn or adapt beyond what they were initially programmed for. Having too many rules can slow them down.
2. **Immutability**
   - By nature, these systems do not change and are unscalable. You can't easily change them or add new rules without a lot of time and expensive complications.
3. **Restricted Intelligence**
   - These systems rely entirely on the rules created by their developers. They can't think for themselves or make decisions outside of those rules.

Machine Learning Approaches advantages

1. Adaptability

Machine learning systems can learn and adjust on their own when they see new data. This helps them make smart decisions and predictions in situations that change quickly.

2. Self-learning

These systems can figure out their own rules by finding patterns in the data. This is useful when dealing with complex tasks that involve many different factors and rules that may change over time.

3. Scalability

Machine learning systems can grow with the needs of a project or environment. They can handle more data, use better resources, and even upgrade to more advanced methods as they go along.

Machine Learning Approaches disadvantages

1. **Depends on Good Training Data**

Machine learning systems need high-quality data to work well. If the data used for training isn't good, the system won't make good decisions. But collecting the right data can be time-consuming and costly.

2. **Complexity**

While machine learning can handle big and complex tasks, it also needs more technical skills to set up. You need a team of experts who can create algorithms, find useful data, and watch over the models.

3. **Limited adaptability**

These systems can adjust and grow, but they cannot adapt to situations or problems outside the realm they are trained for. They can't figure things out on their own or understand things beyond their training. They lack human-like intuition.

2

What are rule-based approach algorithms

2 What are rule-based approaches

In rule-based systems, we have two possible approaches:

1. **Construct explicit rules and patterns**, typically created by human experts or developers who specify how to process and interpret the text

2. Use **lexicons**, i.e., dictionary-based systems that rely on lists of words or phrases with associated sentiment scores

2.1 What are rule-based approaches

Construct explicit rules

- Predefined linguistic rules are used to analyze and process textual data.
- Involves applying a particular set of rules or patterns to capture specific structures, extract information, or perform tasks such as text classification and so on.

The steps:

Rule creation

Depending on the task, linguistic rules are created (grammar rules, semantic rules, or regular expressions)

Rule application

Rules are applied to input data to capture patterns

Rule processing

Text data is processed with the matched rules to extract information

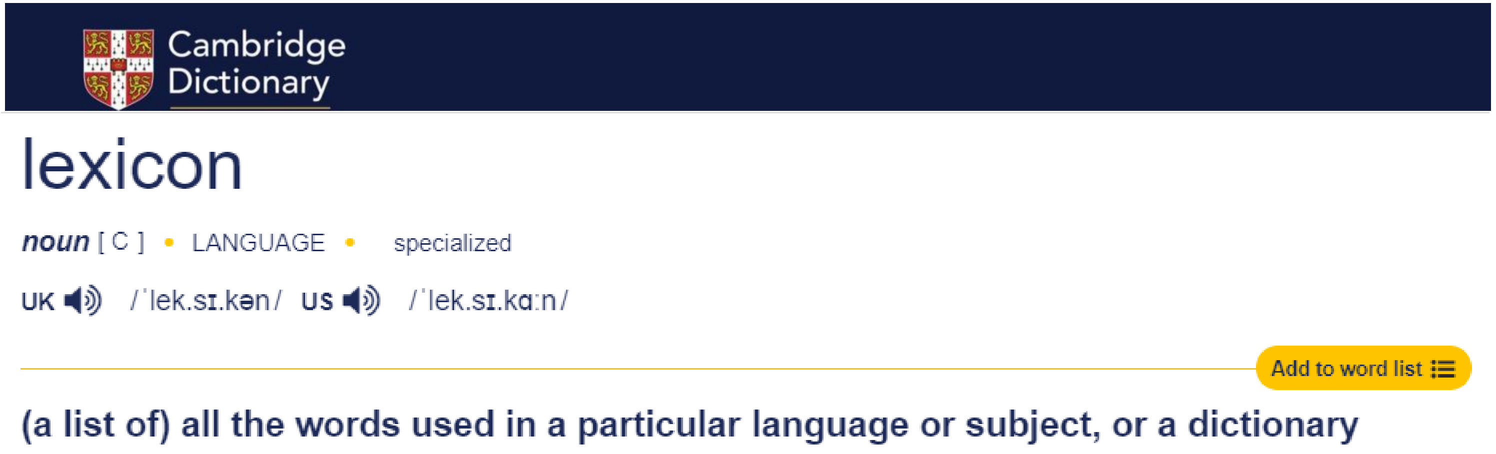
Rule refinement

Rules are iteratively refined by repetitive processing to improve performance

2.2 What are rule-based approaches

Using Lexicons

What is a lexicon?



- A lexicon is like a dictionary that contains a collection of words and has been compiled using expert knowledge.
- It incorporates specific knowledge and has been collected for a specific purpose.

2.2 What are rule-based approaches

Using Lexicons

What is a lexicon?

- We will use sentiment lexicons that contain commonly used words and capture the sentiment associated with them.
  - A simple example of this is the word happy, with a sentiment score of 1, and another is the word frustrated, which would have a score of -1.

- Several standardized lexicons are available for the English language, and the popular ones are **AFINN Lexicon**, **SentiWordNet**, **Bing Liu**'s lexicon, **TextBlob** and **VADER** lexicon, among others.

What are rule-based approaches

Using Lexicons

What is the difference between the lexicons?

- They differ from each other ***in the size of their vocabulary and their representation***.
  - AFINN Lexicon comes in the form of a single dictionary with 3,300 words, with each word assigned a signed sentiment score ranging from -3 to +3. Negative/positive indicates the polarity, and the magnitude indicates the strength.
  - Bing Liu lexicon comes in the form of two lists: one for positive words and another for negative, with a combined vocabulary of 6,800 words.
- Most sentiment lexicons are available in English, but there are also lexicons available in other languages.

We are going to explore VADER and TextBlob lexicons, two of the most used lexicons nowadays

VADER

Valence **A**ware **D**ictionary and s**E**ntiment **R**easoner

VADER

- The first successful rule-based sentiment analysis algorithm developed by Hutto and Gilbert at GA Tech
- Specifically tuned to analyze sentiments in social media
- Many NLP packages implement some form of this algorithm

- It is composed by 7,500 curated lexical features with a proper validated valence score
  - Each feature was rated on a scale from [-1] (Extremely Negative) to [1] (Extremely Positive), where [0] stands for Neutral or Neither

2.2.1 What are rule-based approaches

Using Lexicons - VADER

```
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
analyzer = SentimentIntensityAnalyzer()
sentence = "The food was great!"
vs = analyzer.polarity_scores(sentence)
print(vs)
```

```
{'compound': 0.6588, 'neg': 0.0, 'neu': 0.406, 'pos': 0.594}
```

2.2.1 What are rule-based approaches

Using Lexicons - VADER

negative score **positive score**

```
{'compound': 0.6588, 'neg': 0.0, 'neu': 0.406, 'pos': 0.594}
```

neutral score

The **compound score** will typically fall within the range [-1,1], where:

- A compound score greater than 0 indicates a positive sentiment.
- A compound score less than 0 indicates a negative sentiment.
- A compound score of 0 indicates a neutral sentiment.

Some aspects to take into account when using VADER

Punctuation

Adding an exclamation mark (!) increases the level of intensity without altering the underlying meaning or semantic orientation.

```
vader_scores('Text Mining is great')
vader_scores('Text Mining is great!')
vader_scores('Text Mining is great!!')
vader_scores('Text Mining is great!!!')
```

```
Text Mining is great---------- {'neg': 0.0, 'neu': 0.423, 'pos': 0.577, 'compound': 0.6249}
Text Mining is great!--------- {'neg': 0.0, 'neu': 0.406, 'pos': 0.594, 'compound': 0.6588}
Text Mining is great!!-------- {'neg': 0.0, 'neu': 0.39, 'pos': 0.61, 'compound': 0.6892}
Text Mining is great!!!------- {'neg': 0.0, 'neu': 0.376, 'pos': 0.624, 'compound': 0.7163}
```

Some aspects to take into account when using VADER

Capitalization

Using upper case letters increases the magnitude of the sentiment intensity.

```
vader_scores('Text Mining is great!')
vader_scores('Text Mining is GREAT!')
```

```
Text Mining is great!--------- {'neg': 0.0, 'neu': 0.406, 'pos': 0.594, 'compound': 0.6588}
Text Mining is GREAT!--------- {'neg': 0.0, 'neu': 0.369, 'pos': 0.631, 'compound': 0.729}
```

Some aspects to take into account when using VADER

Degree modifiers

Also called intensifiers, they impact the sentiment intensity by either increasing or decreasing the intensity.

```
vader_scores('Text Mining is great!')
vader_scores('Text Mining is extremely great!')
vader_scores('Text Mining is kind of great!')
```

```
Text Mining is great!--------- {'neg': 0.0, 'neu': 0.406, 'pos': 0.594, 'compound': 0.6588}
Text Mining is extremely great! {'neg': 0.0, 'neu': 0.461, 'pos': 0.539, 'compound': 0.6893}
Text Mining is kind of great!- {'neg': 0.0, 'neu': 0.55, 'pos': 0.45, 'compound': 0.6248}
```

Some aspects to take into account when using VADER

Conjunctions

Use of conjunctions like “but” signals a shift in sentiment polarity, with the sentiment of the text following the conjunction being dominant.

```
vader_scores('Text Mining is great!')
vader_scores('Text Mining is great, but Mondays are terrible')
```

```
Text Mining is great!--------- {'neg': 0.0, 'neu': 0.406, 'pos': 0.594, 'compound': 0.6588}
Text Mining is great, but Mondays are terrible {'neg': 0.327, 'neu': 0.472, 'pos': 0.201, 'compound':
-0.3818}
```

Some aspects to take into account when using VADER

Preceding Tri-gram

By examining the tri-gram preceding a sentiment-laden lexical feature, we catch nearly 90% of cases where negation flips the polarity of the text.

```
vader_scores('Text Mining is great')
vader_scores("Text Mining isn't always that extremely great")
vader_scores("Text Mining isn't always that great")
```

```
Text Mining is great---------- {'neg': 0.0, 'neu': 0.423, 'pos': 0.577, 'compound': 0.6249}
Text Mining isn't always that extremely great {'neg': 0.0, 'neu': 0.577, 'pos': 0.423, 'compound': 0.
659}
Text Mining isn't always that great {'neg': 0.397, 'neu': 0.603, 'pos': 0.0, 'compound': -0.5096}
```

2.2.1 What are rule-based approaches

Using Lexicons - VADER

Some aspects to take into account when using VADER

Handling Emojis, slangs and Emoticons

VADER performs very well with emojis, slangs and acronyms in sentences:

```
vader_scores('I am 😄 today')
vader_scores('😊')
vader_scores('😥')
vader_scores('☹')
```

```
I am 😄 today------------------ {'neg': 0.0, 'neu': 0.522, 'pos': 0.478, 'compound': 0.6705}
😊---------------------------- {'neg': 0.0, 'neu': 0.333, 'pos': 0.667, 'compound': 0.7184}
😥---------------------------- {'neg': 0.275, 'neu': 0.268, 'pos': 0.456, 'compound': 0.3291}
☹---------------------------- {'neg': 0.706, 'neu': 0.294, 'pos': 0.0, 'compound': -0.34}
```

2.2.1 What are rule-based approaches

Using Lexicons - VADER

Some aspects to take into account when using VADER

Slangs

```
vader_scores("Today SUX!")
vader_scores("Today only kinda sux! But BTW I'll get by LOL")
```

```
Today SUX!-------------------- {'neg': 0.779, 'neu': 0.221, 'pos': 0.0, 'compound': -0.5461}
Today only kinda sux! But BTW I'll get by LOL {'neg': 0.109, 'neu': 0.544, 'pos': 0.346, 'compound':
0.6692}
```

Emoticons

```
vader_scores("Make sure you :) or :D today!")
```

```
Make sure you :) or :D today!- {'neg': 0.0, 'neu': 0.294, 'pos': 0.706, 'compound': 0.8633}
```

TEXTBLOB

2.2.2 What are rule-based approaches

Using Lexicons - TEXTBLOB

```
from textblob import TextBlob

sentence = TextBob("The food was great!")
print(sentence.sentiment)
```

```
Sentiment(polarity=1.0, subjectivity=0.75)
```

2.2.2 What are rule-based approaches

Using Lexicons - TEXTBLOB

```
Sentiment(polarity=1.0, subjectivity=0.75)
```

Polarity

- Measures the sentiment or emotional tone of the text.
- By default, it ranges between [-1, 1]
  - -1 indicates a highly negative sentiment
  - 0 indicates a neutral sentiment
  - 1 indicates a highly positive sentiment

Subjectivity

- Measures how objective or subjective the text is.
- By default, it ranges between [0, 1]
  - 0 indicates a highly objective piece of text
    - Fact-based content
  - 1 indicates a highly subjective (opinionated) piece of text
    - Personal opinions, emotions, judgements

What are rule-based approaches

Using Lexicons - TEXTBLOB

The polarity

```
textblob_score('Mondays are marvelous!')
textblob_score('Tuesdays are terrible!')
```

```
Sentiment(polarity=1.0, subjectivity=1.0)
Sentiment(polarity=-1.0, subjectivity=1.0)
```

The subjectivity

```
textblob_score('Text Mining is a wonderful course!')
textblob_score('Text Mining is the process of extracting useful information from unstructured text.')
```

```
Sentiment(polarity=1.0, subjectivity=1.0)
Sentiment(polarity=0.3, subjectivity=0.0)
```

2.2.2 What are rule-based approaches

Using Lexicons - TEXTBLOB

Punctuation

```
textblob_score('Text Mining is great!')
textblob_score('Text Mining is great!!')
textblob_score('Text Mining is GREAT!')
textblob_score('Text Mining is extremely great!')
```

```
Sentiment(polarity=1.0, subjectivity=0.75)
Sentiment(polarity=1.0, subjectivity=0.75)
Sentiment(polarity=1.0, subjectivity=0.75)
Sentiment(polarity=1.0, subjectivity=0.75)
```

2.2.2 What are rule-based approaches

Using Lexicons - TEXTBLOB

Handling Emojis, slangs and Emoticons

```
textblob_score('I am 😄 today')
textblob_score('😊')
textblob_score('😥')
textblob_score('☹')
textblob_score("Today SUX!")
textblob_score("Today only kinda sux! But I'll get by, lol")
```

```
Sentiment(polarity=0.0, subjectivity=0.0)
Sentiment(polarity=0.0, subjectivity=0.0)
Sentiment(polarity=0.0, subjectivity=0.0)
Sentiment(polarity=0.0, subjectivity=0.0)
Sentiment(polarity=0.0, subjectivity=0.0)
Sentiment(polarity=0.4, subjectivity=0.85)
```

What are rule-based approaches

Using Lexicons – VADER vs TEXTBLOB

VADER vs **TEXTBLOB**

VADER

- Use a list of lexical features that are labeled as positive or negative based on their semantic orientation
- Better for **non-formal text** like social media

TEXTBLOB

- Ignores unfamiliar words, and consider the words and phrases that it can assign polarity to average them and get a final score
- Better for **formal text** like books, papers and projects

3

A case study

3 A case study

We are selling a product in Amazon, and we want to understand how people are reacting to our product:

Bluetooth Headphones True Wireless Earbuds

4.1 ★★★★☆ 29,378 ratings | 364 answered questions

Rating 4.1 out of 5, with 29,378 votes!

Seems good, right?

But we have more information than that! So we can go further…

Single button control is an extremely frustrating experience!

Reviewed in the United States on May 6, 2023

Color: Green | Verified Purchase

If you like the feeling of total frustration and lack of control, then these are the earbuds for you.
The sound was good (but, the left bud quit its working audio after 35 days, but controls still function) and they are comfortable, BUT….
The controls are all in ONE BUTTON and dependent on number and length of pushes. It says to read the red/blue lights for confirmation but the buds are in your ears- duh?

Sound and construction comparison to Tozo T5

Reviewed in the United States on August 6, 2023

Color: Black | Verified Purchase

Why I bought these:
I have had a pair of Tozo T5s for about 2.5 years now.. and they sound great, feel great (with tip replacements) and work ok. For about the last half year, the right headphone has become extremely finicky while charging, and often turn themselves on in the case and connect unknowingly to my phone, drain themselves, cause me to miss calls and notifications and are dead by the time I am ready to go mow the lawn or whatever later in the week.

These were on sale for about 1/2 the price last week of the Tozos, had a similar over the ear design a the contacts in the case gave me the impression, they may charge and behave longer. Obviously I haven't had them long enough to do a battery comparison or to see if they randomly connect, I will try to update as I I get enough data.

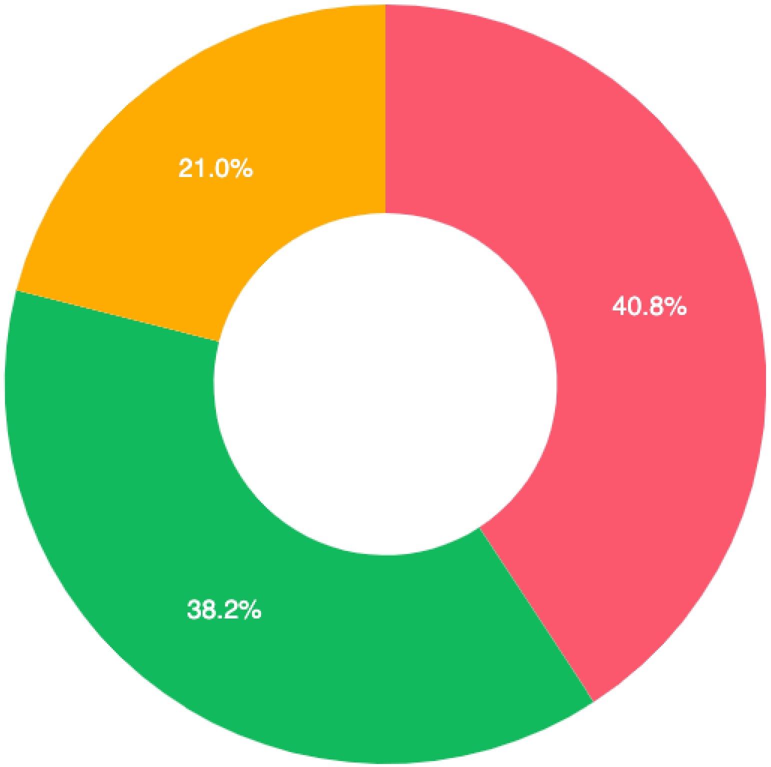
3 A case study

Some possible insights

- Getting the overall sentiment
- Sentiment over time
- Sentiment by rating
- Aspect-based sentiment analysis
- …

3.1 A case study

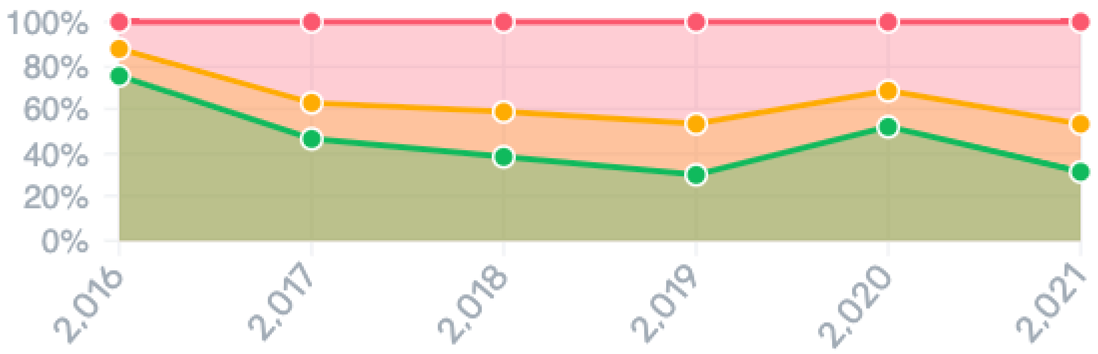
Getting the overall sentiment



You'll notice that these results are very different from Amazon overview.

- That's because we did a more detailed sentiment analysis.
  - We carefully looked at each sentence and even each word in the reviews.
- This way, you get a precise understanding of what customers have written, instead of just counting stars.
  - This analysis can help you find specific issues or problems more accurately and with more details.

Sentiment over time

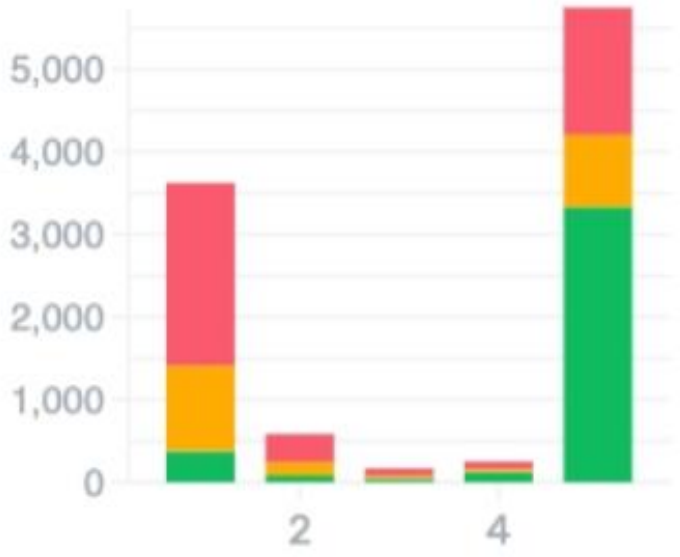


- This chart gives more information about our general sentiment data.
  - It shows how positive, neutral, and negative sentiments in the reviews changed from 2016 to 2021.
- This graph tells us how the content of the reviews evolved over these five years.
  - For example, negative comments decreased from 2019 to 2020 but then went back up to previous levels in 2021.

3.3 A case study

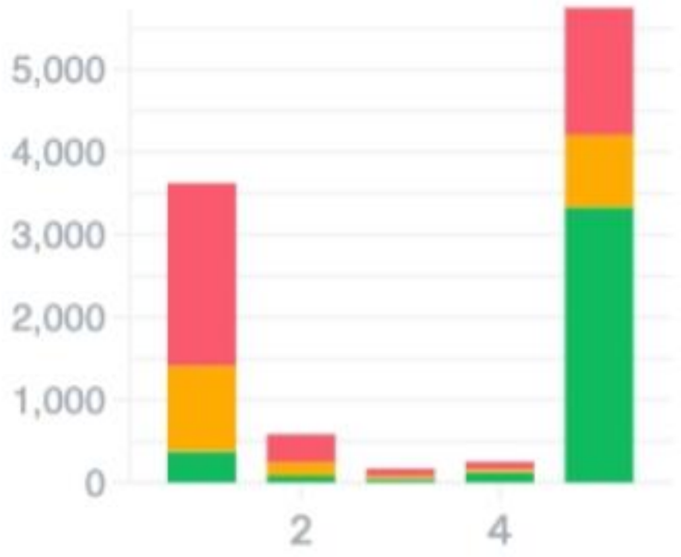
Sentiment by rating

We created the graph by analyzing the written reviews on Amazon, where each category is rated from 1 (Bad) to 5 (Excellent). This breakdown helped us generate the graph you see above.



3.3 A case study

Sentiment by rating

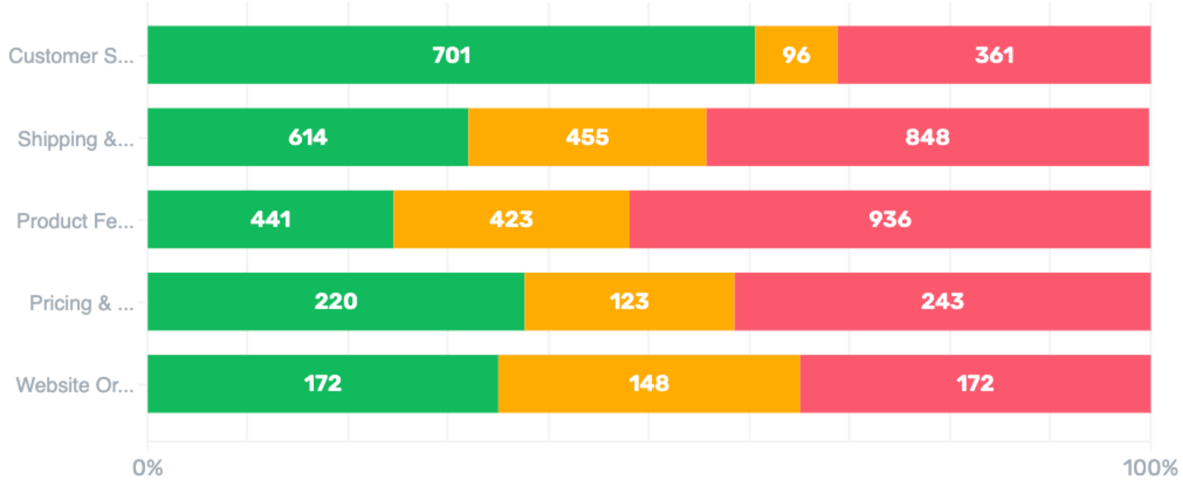


When we look at the results, we can draw a few conclusions:

1. **Amazon's ratings do matter:** Reviews with higher ratings tend to have more positive sentiments, while lower-rated reviews have more negative sentiments. So, the ratings are somewhat reflective of the sentiment in the reviews.
2. **Nuanced reviews:** What's interesting is that all reviews, even the good and bad ones, contain a mix of different sentiments. This shows that our reviews are complex and likely have more valuable insights for us to discover.
3. **Polarized reviews:** Our reviews tend to cluster around the extremes of the rating scale, with many being rated as 5 (Excellent) or 1 (Bad).

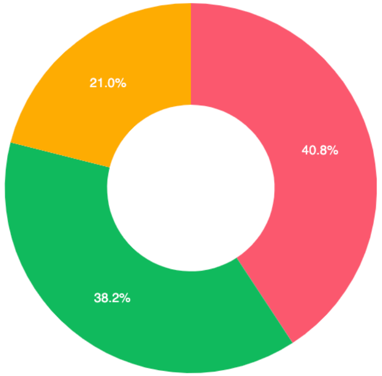
Aspect-based sentiment analysis plot

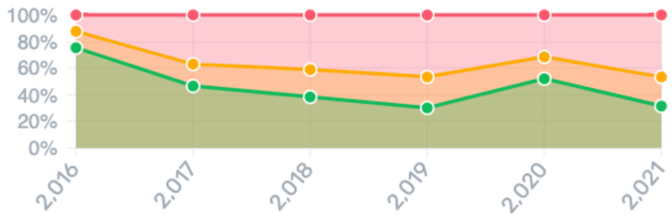
The chart below doesn't just analyze sentiments; it also uses a technique called ***"aspect-based sentiment analysis"*** to connect sentiments with specific features or aspects of a product or service.

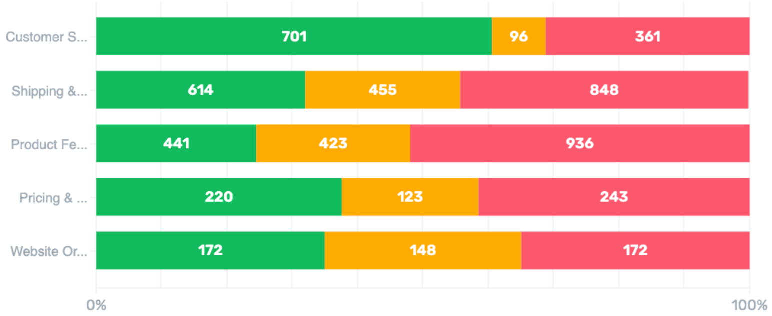


This can be really helpful for us because it allows us to ***pinpoint and address specific problems or concerns***.

3.4 A case study







- These visualizations are excellent starting points to show how valuable sentiment analysis can be.
- However, it's important to note that sentiment analysis has even more capabilities and potential beyond what's shown here.

Practical class…

Sentiment Analysis

Next week…

Spam Filtering

Obrigada!

Morada: Campus de Campolide, 1070-312 Lisboa, Portugal
Tel: +351 213 828 610 | Fax: +351 213 828 611

Acreditações e Certificações da NOVA IMS

Cofinanciado por